A note on the stability of multiclass Markovian queueing networks

Sayee C. Kompalli and Ravi R. Mazumdar

Department of Electrical and Computer Engineering

University of Waterloo, Canada N2L 3G1.

Email: skompall@uwaterloo.ca, mazum@ece.uwaterloo.ca

June 24, 2010

Abstract

In this paper we show that in a multiclass Markovian network with unit rate servers, the condition that the average load ρ at every server is less than unity is indeed sufficient for the stability or positive recurrence for *any* work conserving scheduling policy and *class-independent* routing. We use a variation of the positive recurrence criterion for multidimensional discrete-time Markov chains over countable state spaces due to Rosberg (JAP, Vol. 17, No. 3, 1980) and a monotonicity argument to establish this assertion.

1 Introduction and Overview

In this paper we show that the condition that the load $\rho = \sum_{i=1}^{J_m} \rho_i < 1$ where J_m denotes the number of classes that visit node m at every work conserving unit rate server of a open multiclass Markovian network is sufficient to ensure the stability of the network with class-independent routing. We prove this via a variation of the positive recurrence criterion for multidimensional discrete-time Markov chains over countable state spaces due to Rosberg [1].

Much research effort has focussed on studying the stability of multiclass queueing networks in the last two decades and the amount of the literature thus is understandably vast. Much of the research on this topic was fueled by the example of a queueing network in a deterministic setting that Lu and Kumar studied in [2]. This paper showed the existence of a scheduling policy under which a certain network becomes unstable even though the load on each work conserving server in the network is less than unity, which was counter intuitive given the well known fact that the load is less than unity is a sufficient and necessary condition for the stability of a single queue with multiple classes under *any* work conserving policy. The appeerence of the paper [2] was followed by a number of papers [3, 4, 5, 6, 7, 8, 9, 10, 11] that either established the positive assertion that multiclass networks are stable under specific settings and policies, or proved that under certain specific scheduling policies for which the usual traffic condition $\rho < 1$ at every work conserving server is *not* sufficient to ensure the stability of the network.

Foss and Rybko [11] proposed a condition, called *Jackson-type* condition, under which they showed that $\rho < 1$ at every server is a sufficient condition when a FIFO scheduling

policy used under renewal assumptions on arrival and service processes. The Jackson-type condition roughly means that the routing mechanism is "almost independent" of customer classes. In this paper we also assume that the *routing is class independent* and we show that when arrivals are Poisson and service times are exponential, the network is stable under the usual traffic conditions for *any* work conserving scheduling policy (that could differentiate between classes). In the network models considered in [2], [7], and [6], it should be noted that the customers follow a predetermined fixed route before exiting the network and hence the routing is not class independent.This is different from the situation considered by Kelly [12] where it is shown that multiclass Markovian networks with *class independent* routing are stable under FIFO and PS scheduling under the *usual* condition. This paper essentially extends those results to any work-conserving scheduling policy.

2   The Stability Analysis

A common technique to establish ergodicity of countable space Markov chains is through Foster-Lyapunov approach which consists of finding an appropriate Lyapunov function that satisfies Foster's criterion [13]. Though this approach has proved to be very successful for one dimensional Markov chains, finding such a function proved to be very difficult in the case of multidimensional Markov chains especially networks. Rosberg [1] extended Foster's criterion [13] for $J \geq 2$-dimensional Markov chains by requiring existence of J Lyapunov functions, one for each coordinate of the process. But the applicability of his criterion becomes limited when the state definition for the Markov chain does not have a *fixed* dimension. For example, the well known FCFS scheduling policy requires a variable-length dimension for the state definition. Also, depending on the modeling assumptions, one may only require a fewer component Lyapunov functions than the dimensionality of the underlying state space. Thus, we propose a variation of Rosberg's criterion that expands the scope of the applicability that we use to establish stability for the multiclass network model we consider.

Let $\mathcal{X}$ be a countable set of states over which the irreducible, aperiodic, and discrete-time Markov chain $\{X^n, n \geq 0\}$ takes its values. For any integer $k \geq 1$, define $\{p_{xy}^k, x, y \in \mathcal{X}\}$ to be the k-step transition probability law of the Markov chain $\{X^n, n \geq 0\}$. For any subset $\mathcal{B} \subseteq \mathcal{X}$, we know that the $\lim_{k \to \infty} p_{x\mathcal{B}}^k = \lim_{k \to \infty} \sum_{y \in \mathcal{B}} p_{xy}^k = \pi(\mathcal{B}) \geq 0$ exists and is independent of the initial state x. For any nonnegative-valued function V on $\mathcal{X}$, let us define $\Delta^k V(x) \triangleq \sum_y p_{xy}^k V(y) - V(x)$ to be the k-step drift of the function V in state x. Let $c^* \geq 0$ denote the limit $\lim_{k \to \infty} \frac{\Delta^k V(x)}{k}$, and is independent of x. The following definition was originally proposed in [1].

Definition 2.1 *The sequence $\left\{ \frac{\Delta^k V(x)}{k}, k \geq 1 \right\}$ is said to be uniformly upper bounded (UUB) if, for any $\delta > 0$, there exists a positive integer K such that $\frac{\Delta^k V(x)}{k} < c^* + \delta$ for $k \geq \mathsf{K}$ and $x \in \mathcal{X}$.*

We now propose a variation of the Rosberg's positive criterion [1]. In the rest of the paper, we will use the same notation $\mathcal{X}$ to denote both the state space and its subsets. The distinction is made through the usage of subscripts, i.e., $\mathcal{X}_j$ denotes a subset.

Theorem 2.1 (A variation of Rosberg's criterion [1]) *Let $J \geq 2$ be an integer.*

Assumption 2.1 *There exists a collection* $\mathcal{P} = \{\mathcal{P}_1, \mathcal{P}_2, \ldots, \mathcal{P}_J\}$ *of partitions of the set* $\mathcal{X}$ *where* $\mathcal{P}_j = \{\mathcal{X}_j, \mathcal{X}_j^c\}$, *and nonnegative-valued functions* $\{V_j(x), x \in \mathcal{X}\}$ *for* $1 \leq j \leq J$ *such that the drift* $\Delta V_j(x)$ *of the function* V_j *in the state* x *has the following form:*

$$\Delta V_j(x) \;\; \leq \;\; \begin{cases} \eta_j & \text{for } x \in \mathcal{X} \\ -\epsilon_j & \text{for } x \in \mathcal{X}_j^c, \end{cases} \tag{1}$$

where $\epsilon_j > 0$ *and* $\eta_j \geq 0$.

Assumption 2.2 *There exist partitions* $\{\mathcal{A}_{j,k}, \mathcal{A}_{j,k}^c\}$, $k \geq 1$ *and* $1 \leq j \leq J$, *of the set* $\mathcal{X}$ *with the following two properties:*

(i) $p_{xy}^l = 0$, $0 \leq l \leq k-1$, *for* $x \in \mathcal{A}_{j,k}^c$ *and* $y \in \mathcal{X}_j$

(ii) $\cap_j \mathcal{A}_{j,k}$ *is a finite set*

Assumption 2.3 *For* $1 \leq j \leq J$, *the sequence* $\left\{\frac{\Delta^k V_j(x)}{k}, k \geq 1\right\}$ *is uniformly upper bounded (UUB).*

Then the Markov chain $\{X_n, n \geq 1\}$ *is positive recurrent.* ∎

Before we provide a proof of Theorem 2.1 we first discuss how our variation is different from the original theorem in [1].

In [1] it is assumed that the the state space of the MC model denoted by $\mathcal{X} = \mathbb{Z}_+^J$, the J-dimensional non-negative integer space, for some integer $J \geq 2$ and an *equal number* J partitions denoted by $\{\mathcal{X}_j, \mathcal{X}_j^c\}$, $1 \leq j \leq J$, of the countable space $\mathcal{X}$, and also the same number J of Lyapunov functions $\{V_j(x), x \in \mathcal{X}\}$, $1 \leq j \leq J$. In the variation proposed in this paper, *we do not require the countable space* $\mathcal{X}$ *to have a fixed predetermined dimension.* This corresponds to the Assumption 2.1 of Theorem 2.1. Hence we are free to choose an appropriate number of Lyapunov functions and the corresponding suitable partitions of the state space $\mathcal{X}$. We believe this generalization is useful in many situations of interest when the state space is not of fixed predetermined dimension.

For instance, when the FIFO scheduling policy is implemented at a server the state of each queue is characterized not just by the total number of customers present in the queue, but also by the class of the customers and the relative order of arrival of the customers of different classes. Thus the state description of the server is *not* of fixed dimension but of variable dimension, and potentially can grow very large depending on the number of customers.

Proof: Since the proof of Theorem 2.1 is almost identical to the Proof of Theorem 2 in [1] we provide here only the main argument, omitting many intermediate results and supporting arguments which can be found in [1].

Let us fix an arbitrary $\delta > 0$. From Assumption 2.3 of Theorem 2.1, it follows that there exits a positive integer K such that $\frac{\Delta^k V_j(x)}{k} \leq c_j^* + \delta$ for $k \geq \mathsf{K}$ and $1 \leq j \leq J$. Let us pick one such K, and then introduce the set of functions $\{g_j^\mathsf{K}(x); x \in \mathcal{X}\}$ such that the following holds:

$$\Delta^\mathsf{K} V_j(x) \;\; = \;\; -g_j^\mathsf{K}(x) + \mathsf{K}\left(c_j^* + \delta\right) \tag{2}$$

Two observations on the functions g_j^K are in order: the first and the obvious observation is that $g_j^\mathsf{K}(x) \geq 0$ for $x \in \mathcal{X}$. Also, since $\Delta^\mathsf{K} V_j(x) \leq -\mathsf{K}\epsilon_j$ for $x \in \mathcal{A}_{j,\mathsf{K}}^c$ (Lemma 2 of [1]), we

have that $g_j^{\mathsf{K}}(x) \geq \left(c_j^* + \delta + \epsilon_j\right)$ for $x \in \mathcal{A}_{j,\mathsf{K}}^c$. Set $\epsilon = \min_j \epsilon_j$ and $\delta = \min_j \delta_j$. As a result, we have the obvious deduction that $\max_j g_j^{\mathsf{K}}(x) \geq \min_j \mathsf{K}\left(c_j^* + \delta + \epsilon_j\right) = (c^* + \delta + \epsilon)$ for $x \in \cup_j \mathcal{A}_{j,\mathsf{K}}^c$. Hence $\max_j g_j^{\mathsf{K}}(x) < \min_j \mathsf{K}\left(c^* + \delta + \epsilon\right)$ implies that $x \in \cap_j \mathcal{A}_{j,\mathsf{K}}$. We should note that $x \in \cap_j \mathcal{A}_{j,k}$ *need not imply that* $\max_j g_j^{\mathsf{K}}(x) < \mathsf{K}\left(c^* + \delta + \epsilon\right)$.

Denote by $\mathsf{E}_x\left(g_j^{\mathsf{K}}(X^n)\right)$ the expectation of $g_j^{\mathsf{K}}(X^n)$ given that $X^0 = x$ and by $p_x(X^n \in A)$ the probability that $X^n \in A$ given that $X^0 = x$. Now

$$
\begin{aligned}
\frac{\Delta^{n\mathsf{K}} V_j(x)}{n} &= \sum_{y \in \mathcal{X}} \frac{1}{n} \sum_{l=0}^{n-1} p_{xy}^{l\mathsf{K}} \Delta^{\mathsf{K}} V_j(y) \\
&\overset{(a)}{=} \sum_{y \in \mathcal{X}} \frac{1}{n} \sum_{l=0}^{n-1} p_{xy}^{l\mathsf{K}} \left[-g_j^{\mathsf{K}}(y) + \mathsf{K}(c_j^* + \delta) \right] \\
&= -\sum_{y \in \mathcal{X}} \frac{1}{n} \sum_{l=0}^{n-1} p_{xy}^{l\mathsf{K}} g_j^{\mathsf{K}}(y) + \mathsf{K}(c_j^* + \delta) \\
&= -\frac{1}{n} \sum_{l=0}^{n-1} \mathsf{E}_x\left(g_j^{\mathsf{K}}\left(X^{l\mathsf{K}}\right)\right) + \mathsf{K}(c_j^* + \delta)
\end{aligned}
$$

where (a) follows from (2).

Since $\lim_{n\to\infty} \frac{\Delta^{n\mathsf{K}} V_j(x)}{n\mathsf{K}} = c_j^*$, we have that $\frac{1}{n}\sum_{l=0}^{n-1} \mathsf{E}_x\left(g_j^{\mathsf{K}}\left(X^{l\mathsf{K}}\right)\right) = \mathsf{K}\delta$. Now

$$
\begin{aligned}
\liminf_{n\to\infty} \frac{1}{n} p_x\left(\max_j g_j^{\mathsf{K}}(x) < \mathsf{K}\left(c^* + \delta + \epsilon\right)\right) &\overset{(b)}{\geq} 1 - \limsup_{n\to\infty} \frac{1}{n} \sum_{l=0}^{n-1} \sum_{j=1}^{J} \frac{\mathsf{E}_x\left(g_j^{\mathsf{K}}\left(X^{l\mathsf{K}}\right)\right)}{\mathsf{K}\left(c^* + \delta + \epsilon\right)} \\
&\geq 1 - \frac{1}{\mathsf{K}\left(c^* + \delta + \epsilon\right)} \times \\
&\qquad \sum_{j=1}^{J} \limsup_{n\to\infty} \frac{1}{n} \sum_{l=0}^{n-1} \mathsf{E}_x\left(g_j^{\mathsf{K}}\left(X^{l\mathsf{K}}\right)\right) \\
&= 1 - \frac{J\mathsf{K}\delta}{\mathsf{K}\left(c^* + \delta + \epsilon\right)} \\
&= 1 - \frac{J\delta}{\left(c^* + \delta + \epsilon\right)},
\end{aligned}
$$

where (b) follows from Lemma 1 in [1].

We note that there exists a $\delta_0 > 0$ such that $1 - \frac{J\delta}{(c^* + \delta_0 + \epsilon)} > 0$. Define the set

$$
\mathcal{A}_0 = \left\{ x \in \mathcal{X} : \max_j g_j^{\mathsf{K}}(x) < \mathsf{K}\left(c^* + \delta_0 + \epsilon\right) \right\}
$$

We can observe that $\mathcal{A}_0 \subseteq \cap_j \mathcal{A}_{j,k}$ is a finite set. Hence it follows that for the finite set $\mathcal{A}_0$,

$$
\liminf_{n\to\infty} \frac{1}{n} \sum_{l=0}^{n-1} p_x\left(X^{l\mathsf{K}} \in \mathcal{A}_0\right) > 0
$$

Since the chain is assumed to be irreducible and aperiodic, it follows that the Markov chain is positive recurrent. ∎

3 Multiclass Network Model

Consider an open queueing network consisting of J work conserving single-servers, each with an infinite capacity queue. Let $\alpha = 1, 2, \ldots$ denote the class of an arriving customer and customers of class α arrive at server $j, j = 1, 2, \ldots, J$ as a Poisson process with rate $\lambda_{\alpha,j}$ and a customer of class α at node j requires a service time that is exponentially distributed with rate $\mu_{\alpha,j}$. After completing service a customer of class α is routed to server k with probability $r_{j,k}$ that is the *same* for all classes. It is assumed that the scheduling policy at any server is a stationary, *work-conserving* policy.

One way of constructing this network is to consider a model in which customers from outside to the network arrive in a Poisson process of rate λ. On arrival, the customer joins the queue attached to server j as a class α customer with the probability $q_{\alpha,j}$ such that $\sum_{\alpha,j} q_{\alpha,j} = 1$. Thus $\lambda_{\alpha,j} = \lambda q_{\alpha,j}$. Once assigned, a customer will retain its class till it exits the network. Let $R = [r_{j,k}]$ denote the routing probability matrix and by construction R^T is *invertible* since by assumption the network is open.

Let $\{\Lambda_{\alpha,j}\}$, where $\Lambda_{\alpha,j}$ is the equilibrium rate at which customers of the class α arrive to the jth server, be the solution of the following set of equations

$$\Lambda_{\alpha,j} = \lambda q_{\alpha,j} + \sum_{k=1}^{J} \Lambda_{\alpha,k}\, r_{k,j} \tag{3}$$

written for each class α and for each server j.

Let us define $\rho_{\alpha,j}$ as the load on server j from customers of the class α, then

$$\rho_{\alpha,j} = \frac{\Lambda_{\alpha,j}}{\mu_{\alpha,j}}$$

Let $\Gamma_{\alpha,j}^{k}$ denote the expected number of times a customer of the class α, who is presently in the queue j, visits the server k before it departs the network. Then $\Gamma_{\alpha,j}^{k}$ satisfies the following equations:

$$\left.\begin{aligned}
\Gamma_{\alpha,j}^{k} &= \delta_{j,k} + \sum_{l=1}^{J} r_{j,l}\Gamma_{\alpha,l}^{k} \\
\sum_{\alpha}\sum_{k} \lambda q_{\alpha,k}\Gamma_{\alpha,k}^{j} &= \sum_{\alpha} \lambda_{\alpha,j}
\end{aligned}\right\} \tag{4}$$

where $\delta_{j,k}$ is the Kronecker Delta function. Let $\mathcal{X}$ denote the countable state space for the Markov chain $\{X_n, n \geq 1\}$, modeling the evolution of the numbers of customers of different classes in various queues. We assume that the state definition is detailed enough that in every state x we can obtain the class and positional information of every customer in every queue. Let $\{p_{x,y}, x, y \in \mathcal{X}\}$ denote the transition probabilities of the Markov chain.

In the following we show that for a given multiclass network one can find a multiclass network but with *single customer service rate* (i.e., $\mu_{\alpha,j} = \mu$) such that stability of the *single customer service rate* network implies stability of the actual multiclass network. We will see that this reduction of multi service rates to single service rate coupled with the class-independent routing assumption helps us easily establish some monotonicity properties (Lemma 3.1) in the single service rate network.

Let $\mathcal{S}$ represent a network in which the Poisson arrival rate λ and the service rates $\{\mu_{\alpha,j}\}$ are such that $\sum_{\alpha} \rho_{\alpha,j} < 1$ for each server j. We now construct a network $\mathcal{S}'$

which will be identical to the actual network $\mathcal{S}$ except for the Poisson arrival rate and average service times of the customers. Specifically, we will choose a λ' and μ such that $\rho_{\alpha,j} < \frac{\Lambda'_{\alpha,j}}{\mu} < \rho_{\alpha,j} + \eta_{\alpha,j} < 1$ and $\sum_\alpha \rho'_{\alpha,j} = \frac{1}{\mu} \sum_\alpha \Lambda'_{\alpha,j} < 1$ hold for all classes α and all servers j where $\eta_{\alpha,j}$ are arbitrarily small positive numbers. We note here that $\Lambda'_{\alpha,j}$ solve the traffic equation (3) with λ' in place of λ. Equivalently

$$\max_{\alpha,j} \frac{\Lambda'_{\alpha,j}}{\rho_{\alpha,j} + \eta_{\alpha,j}} < \mu < \min_{\alpha,j} \mu_{\alpha,j}, \quad \text{and} \quad \mu > \max_j \sum_\alpha \Lambda'_{\alpha,j}$$

should hold. This becomes feasible when λ' is chosen sufficiently small. Thus, we have constructed a network $\mathcal{S}'$ in which customers arrive at a lower rate and spend longer times on an average in each of the servers they visit. Furthermore, the load due to class α at the server j in $\mathcal{S}'$ is larger than the corresponding in the actual network $\mathcal{S}$. We remark here that even though customers in the network $\mathcal{S}'$ have the same service rate μ they are still identifiable by their class.

Following the uniformization technique [14], we will now obtain the discrete-time Markov chain queueing models $\{X_n, n \geq 0\}$ and $\{Y_n, n \geq 0\}$ of the multiclass networks $\mathcal{S}$ and $\mathcal{S}'$, respectively. For the purposes of uniformization, let us choose two numbers Q_1 and Q_2 such that $Q_1 > \lambda + J \max_{\alpha,j} \mu_{\alpha,j}$ and $Q_2 > \lambda' + J\mu$ and $\frac{\lambda}{Q_1} = \frac{\lambda'}{Q_2}$ where Q_1 and Q_2 are the respective Poisson rates that achieve uniformization for the networks $\mathcal{S}$ and $\mathcal{S}'$, respectively. Note that Q_1 and Q_2 are so chosen that at every time epoch the probability that a customer arrives is *same* in both the chains.

With this construction we now have two $\{X_n\}$ and $\{Y_n\}$ that are the same except for the probability for service completion at a given time epoch: *all servers in the network $\mathcal{S}'$ are slower than the corresponding servers in the network $\mathcal{S}$.*

Next we demonstrate a construction in which the Markov chains $\{X_n, n \geq 0\}$ and $\{Y_n, n \geq 0\}$ will be defined on a common probability space and then we show that the total number of customers in the network $\mathcal{S}'$ will always be at least as large as the total number of customers in the network $\mathcal{S}$ under the condition that both networks are initialized in the *same state*. Let $\{A_n, n \geq 1\}$ be a sequence of i.i.d. random variables with the distribution $p(A_n = 1) = \frac{\lambda}{T_1} = \frac{\lambda'}{T_2} = 1 - p(A_n = 0)$ where $A_n = 1$ denotes an arrival at the nth epoch in *both* the networks. For $1 \leq j \leq J$, let $\{B_{j,k}, k \geq 1\}$ be a sequence of i.i.d. random variables with the distribution $p(B_{j,k} = l) = r_{j,l}$ where $B_{j,k}$ denotes the server to be chosen by the kth departure from the jth server in *both* the networks. Also, by suitable construction we can generate departure events so that a when the jth server is non-empty in $\mathcal{S}$ as well as $\mathcal{S}'$ then a departure from the jth server in $\mathcal{S}'$ implies a departure from the jth server in $\mathcal{S}$. An implication of this construction is that the departure epoch of kth customer from the jth server in $\mathcal{S}'$ will be *no sooner* than the corresponding departure in the network $\mathcal{S}$. Thus the total number of customers in the network $\mathcal{S}'$ will always be at least as large as the total number of customers in the network $\mathcal{S}$. Thus stability of the network $\mathcal{S}'$ implies stability of the actual network $\mathcal{S}$.

Following the arguments above, from now on we concentrate exclusively on the network with a single customer service rate.

Theorem 3.1 *The condition $\sum_\alpha \rho_{\alpha,j} < 1$ for $1 \leq j \leq J$ is a sufficient condition for positive recurrence of the Markov chain $\{X_n, n \geq 1\}$.* ∎

Before we prove Theorem 3.1, we need to establish the following Lemma.

Lemma 3.1 *Let $x \in \mathcal{X}$ be a non-zero state and $p_{x,\mathcal{X}_j}^n = \sum_{y \in \mathcal{X}_j} p_{x,y}^n$. Then $p_{x,\mathcal{X}_j}^n \leq p_{0,\mathcal{X}_j}^n$ under class-independent routing.* ∎

Proof: Let the notation X_n^x, $n \geq 1$, represent the Markov chain whose state at time $n = 0$ is x. Let $y_j^n(x) = \sum_\alpha x_{\alpha,j}^n$ denote the total number of customers in the jth queue after n transitions when the state at $n = 0$ is x. Let now x be a non-zero state and 0, the zero state. We need to show that $y_j^n(x)$ is *stochastically larger* than $y_j^n(0)$; in notation $y_j^n(x) \geq_{\text{s.t.}} y_j^n(0)$. But this implies that $p_{x,\mathcal{X}_j}^n \leq p_{0,\mathcal{X}_j}^n$ for $n \geq 1$. Following Strassen's theorem [15] it is enough to show the same on a *common probability space* formed by the customer arrival process, random variables modeling service times of customers of different classes at different servers, and also the random variables that model routing decisions of customers at each of the J servers.

Because of the assumption that customers routing decisions are independent of their class, it is not hard to see that $y_j^n(x) \geq_{\text{s.t.}} y_j^n(0)$ and hence $p_{x,\mathcal{X}_j}^n \leq p_{0,\mathcal{X}_j}^n$, $n \geq 1$ given that x is a non-zero state. That is, the total number of customers in each queue in a non-zero state always dominate the corresponding quantities in a smaller state. ∎

Proof of Theorem 3.1: By letting $x_{\alpha,k}$ denote the number of customers of the class α at the server k, we consider the Lyapunov functions $V_j(x)$, one for each server j, that were originally proposed in [1].

$$V_j(x) = \sum_\alpha \sum_{k=1}^{J} x_{\alpha,k} \Gamma_{\alpha,k}^j$$

The Lyapunov function $V_j(x)$ can be interpreted as the *virtual work load* on the server j in state x, whereas $x_j = \sum_\alpha x_{\alpha,j}$, the total number of customers who are presently in the queue j, denotes the *physical work load* on the server j. We now define the sets $\mathcal{X}_j = \{x \in \mathcal{X} : x_j = 0\}$ and then $\mathcal{X}_j^c = \{x \in \mathcal{X} : x_j \geq 1\}$ so that $\{\mathcal{X}_j, \mathcal{X}_j^c\}$ is a partition of the state space $\mathcal{X}$. Define $Q = \lambda + J\mu$. Our next task is establish the drift

$$\Delta V_j(x) = \begin{cases} \frac{1}{Q}\left(\sum_\alpha \Lambda_{\alpha,j} - \mu\right) & \text{if } x \in \mathcal{X}_j^c \\ \frac{1}{Q}\sum_\alpha \Lambda_{\alpha,j} & \text{if } x \in \mathcal{X}_j \end{cases}$$

We identify the following events in the system that result in a state transition:

E1 An exogenous customer arrival of class α to the server j with probability $\frac{\lambda_{\alpha,j}}{Q}$,

E2 A departure from the jth server that moves to the kth server with probability $\frac{\mu r_{j,k}}{Q}$,

E3 A departure from the jth server that exits the network with probability $\frac{\mu(1 - \sum_{k=1}^{J} r_{j,k})}{Q}$ and,

E4 A transition that takes the chain back to the same state with probability $1 - \frac{\lambda}{Q} - \sum_{j:x_j \neq 0} \frac{\mu}{Q}$.

Then we can derive the drift $\Delta V_j(x)$ as

$$Q\Delta V_j(x) = \underbrace{\sum_\alpha \sum_{k=1}^{J} \Gamma_{\alpha,k}^j \lambda q_{\alpha,k}}_{E1} + \underbrace{\sum_\alpha \sum_{m,n} \left(\Gamma_{\alpha,n}^j - \Gamma_{\alpha,m}^j\right)\mu r_{m,n}}_{E2} + \underbrace{\sum_\alpha \sum_m -\Gamma_{\alpha,m}^j \mu[1 - \sum_{k=1}^{J} r_{m,k}]}_{E3}$$

$$\overset{(*)}{=} \sum_\alpha \Lambda_{\alpha,j} - \mu\mathbb{I}\{x_j > 0\}$$

7

where $(*)$ follows from the equations (4).

To verify the UUB condition (Assumption 2.3 of Theorem 2.1), we first note that

$$
\begin{aligned}
Q\frac{\Delta V_j^n(x)}{n} &= \sum_{y \in \mathcal{X}_j} Q\Delta V_j(y)\frac{1}{n}\sum_{k=0}^{n-1} p_{x,y}^k + \sum_{y \in \mathcal{X}_j^c} Q\Delta V_j(y)\frac{1}{n}\sum_{k=0}^{n-1} p_{x,y}^k \\
&= \left(\sum_\alpha \Lambda_{\alpha,j}\right)\frac{1}{n}\sum_{k=0}^{n-1} p_{x,\mathcal{X}_j}^k + \left(\sum_\alpha \Lambda_{\alpha,j} - \mu\right)\frac{1}{n}\sum_{k=0}^{n-1} p_{x,\mathcal{X}_j^c}^k \\
&= \sum_\alpha \Lambda_{\alpha,j} - \mu\frac{1}{n}\sum_{k=0}^{n-1} p_{x,\mathcal{X}_j^c}^k
\end{aligned}
$$

From Lemma 3.1 this implies that $\frac{1}{n}\Delta V_j^n(x) \leq \frac{1}{n}\Delta V_j^n(0)$, thus proving the UUB condition. ∎

References

[1] Z. Rosberg. "A Positive Recurrence Criterion Associated with Multidimensional Queueing Processes". *Journal of Applied Probability*, 17(3):790–801, 1980.

[2] S. H. Lu and P. R Kumar. Distributed scheduling based on due dates and buffer priorities. *IEEE Trans. Automat. Control*, 36:1406–1416, 1991.

[3] P. R. Kumar and T. I Seidman. Dynamic instabilities and stabilization methods in distributed real-time scheduling of manufacturing systems. *IEEE Trans. Automat. Control*, 35:289–298, 1990.

[4] P. R Kumar and S Meyn. Stability of Queueing Networks and Scheduling Policies. *IEEE Trans. Automat. Control*, 40(2):251–260, February 1995.

[5] A. N. Rybko and A. L Stolyar. On the ergodicity of random processes that describe the functioning of open queueing networks. *Problemy Peredachi Informatsii*, 28(3-26), 1992.

[6] M Bramson. Instability of FIFO queueing networks. *Ann. Appl. Probab*, 4:414–431, 1994.

[7] T. I Seidman. First come, first served can be unstable! *IEEE Trans. Automat. Control*, 39, 1994.

[8] J. G. Dai. On positive harris recurrence of multiclass queueing networks: A unified approach via fluid limit models. *Ann. Appl. Probab.*, 5:49–77, 1995.

[9] A. L Stolyar. On the stability of multiclass queueing networks: A relaxed sufficient condition via limiting fluid processes. *Markov Process. Related Fields*, 1:491–512, 1995.

[10] M Bramson. Convergence to equilibria for fluid models of fifo queueing networks. *Queueing Systems Theory Appl.*, 22:5–45, 1996.

[11] S. Foss and A. Rybko. Stability of Multiclass Jackson-Type Networks. *Markov Processes and Related Fields*, 2(3):461–487, 1996.

[12] F. P. Kelly. *Reversibility and Stochastic Networks.* John Wiley & Sons Ltd, 1979.

[13] F. G. Foster. "On the stochastic matrices associated with certain queueing processes". *Ann. Math. Statist*, 24:355–360, 1953.

[14] S.A Lippman. Applying a new devise in optimization of exponential queueing systems. *Oper. Res.*, 23:687–710, 1975.

[15] F. Baccelli and P. Bremaud. *Elements of Queueing Theory*, volume 26 of *Applications of Mathematics*. Springer, second edition.